

PRATHAMESH KHOLE

Machine Learning Researcher | Ph.D. Student

📍 Santa Cruz, CA 📞 831-295-9354 ✉ pkhole@ucsc.edu [in](https://www.linkedin.com/in/PrathameshKhole) Prathamesh Khole [github](https://github.com/prathameshkhole) prathameshkhole.github.io

Education

University of California Santa Cruz

Ph.D. in Computer Science and Engineering

June 2029

Santa Cruz, California

University of California Santa Cruz

Master of Science in Computer Science and Engineering

August 2024

Santa Cruz, California

Pune Institute of Computer Technology (University of Pune)

Bachelor of Engineering in Computer Science

May 2021

Pune, India

Relevant Coursework

- | | | | |
|------------------------------------|----------------------------------|--------------------------|-------------------------|
| • Advanced Machine Learning | • Artificial Intelligence | • Deep Learning | • Data Analytics |
| • Distributed Systems | • Machine Learning | • Data Structures | • Operating Systems |
| | | • Analysis of Algorithms | • Robotics |

Experience

University of California Santa Cruz

Graduate Research Assistant — Machine Learning | Medical Imaging | Computer Graphics

January 2023 – Present

Santa Cruz, California

- Working under guidance of Professor **Razvan Marinescu** as a part of his lab at **UCSC**.
- Worked with **Python** frameworks including **PyTorch** and **Numba** to reverse diffusion Magnetic Resonance Imaging (MRI) simulation.
- Developed and implemented **physics based biomedical image simulators** replicating state of the art models in performance.
- Utilized medical image visualization tools like **Free Surfer**, **ITK Snap**, and **Paraview** to better understand and visualize results as well as fix implementation issues.

University of California Santa Cruz

Teaching Assistant — Unix | Systems | C programming

January 2025 – Present

Santa Cruz, California

- **Course Instruction:** Conducted discussions sessions for **CSE-13S Systems and C programming** ensuring comprehensive coverage of key concepts. Including data structures, syntax and shell scripting

University of California Santa Cruz

Teaching Assistant — Logic Design | Assembly language

January 2023 – June 2024

Santa Cruz, California

- **Course Instruction:** Conducted lab sessions for **CSE-12 Computer Systems and Assembly Language** ensuring comprehensive coverage of key concepts. Demonstrated the design and creation of intricate circuits using **Digital Logic** and **Boolean Algebra** principles.
- **RISC-V Assembly Instruction:** Delivered in-depth instruction on **RISC-V Assembly Language**, equipping students with the knowledge and skills to navigate and apply programming fundamentals to assembly language effectively.

ZS Associates

Data Engineer — Data Analysis | Data Pre-processing

August 2021 – June 2022

Pune, India

- **Efficient Data Preprocessing:** Leveraged **Python**, **PySpark**, and **HiveQL** to revolutionize the preprocessing workflow, **achieving over 90% time reduction**, ensuring data is correctly prepared for machine learning applications.
- **Enhanced Product Penetration:** Utilized advanced data analytics and refined **SQL** queries to derive insights into sales trends, enabling strategic expansion into untapped sales territories.
- **Enhanced Customer Engagement:** Through meticulous data analysis and refining **SQL** queries in workflow, crafted more targeted recommendation algorithms, resulting in an approximate **10% improvement in suggestion reach**.
- **Data Integration and Extraction:** Utilized **SQL** and **HiveQL** for efficient data extraction in production, ensuring seamless integration into the machine learning model.
- **Sales & Customer Engagement Dashboard:** Designed and integrated a dynamic dashboard to visualize sales and customer engagement metrics, providing actionable insights for client decision-making and strategy optimization.

Schneider Electric Systems Middle East

Project Intern

June 2019 – August 2019

Al-Ahmadi, Kuwait

- **Process Graphics Validation:** Conducted comprehensive testing initiatives to ensure Process Graphics functionality was in perfect alignment with project specifications and requirements.
- **Database Integrity Assurance:** Performed validations of the database, ensuring the Database configuration adhered strictly to input specifications, guaranteeing data accuracy and reliability.

Projects

Differentiable Diffusion Magnetic Resonance Imaging

March 2023 – Present

University of California Santa Cruz

- Creating a framework to make the process of acquiring **diffusion MRI signals** from a given shape or mesh of brain **differentiable**.
- The framework would also be able to reverse the **diffusion MRI** process, such that we can obtain the mesh or **3D brain structure** given a **diffusion MRI signal**.
- Implemented a **physics based simulator** to simulate the core process of diffusion MRI signal acquisition in **Python** for any given mesh, replicating state of the art models.

Octo-Wumpus Protocol for Fair Scheduling

March 2024

University of California Santa Cruz

- Designed an enhanced lottery scheduling algorithm to ensure fairness using a **queue-based scheduler** and **alpha-inflation** for dynamically boosting starved processes.
- Implemented in **Python** with multithreading to test on parallel tasks like merge sort, DFS, and file operations, achieving fairness without sacrificing efficiency.
- Improved fairness and reduced starvation in scheduling while maintaining the probabilistic nature of lottery algorithms.

Image Classification using Transfer Learning

December 2023

University of California Santa Cruz

- Fine-tuned **Swin Transformers** (`swin_base_patch4_window7_224`, `swin_large_patch4_window7_224`) from the **timm library** to classify 1000 images across 100 classes, despite the challenge of having only 10 examples per class, achieving up to **73.7%** accuracy.
- Addressed overfitting caused by the limited dataset using **data augmentation** techniques such as CutMix, MixUp, Random Erase, and Random Crop; optimized using **SGD** and **Cosine Annealing**.
- Implemented in **Python** with **PyTorch**, leveraging advanced learning rate scheduling and balanced dataset splits to overcome data constraints and ensure robust performance.

Red-Black Tree Based Oblivious Random Access Machine

March 2023

University of California Santa Cruz

- Developed an **Oblivious Random Access Machine (ORAM)**, based on Path ORAM which conceals the users access pattern.
- The ORAM works by using a Red-Black tree as the logical tree in the background for faster information access of the stored data.
- The concealing of access patterns is done by performing a series of dummy reads for every access to data (read or write), so that the overall access pattern appears uniform or same for all operations to the observer.
- The implementation is slightly faster than previous implementations for deletions as Red-Black trees are faster than AVL trees for deletions, implemented in **C++** for speed.

Philanthropy on Blockchain (Ethereum Based DApp)

May 2021

Pune Institute of Computer Technology (University of Pune)

- Engineered a **Ethereum blockchain** based **decentralized application** for donation management, aimed at transparency and vote based approval for donations.
- Ensuring a secure, trackable donation platform (pictures, receipts) without third-party involvement, with an option to rollback donations if trust voting falls below a threshold.
- Implemented using **Ropsten Ethereum Test network**, **Solidity** for writing and implementing smart contracts, using **nextJS**, **ReactJS** and **GraphJS** for frontend.

Document Reader & Extractor (Image Processing and OCR)

April 2020

Pune Institute of Computer Technology (University of Pune)

- Developed a tool in **Python** using **image processing** libraries like **CV2** and **PIL**.
- Integrated with **PyTesseract** for **OCR** process to **extract text from documents**, and used **PIL** and **Pandas** to save the extracted text as csv or text document.
- Developed with a goal to allow for faster processing of handwritten or printed documents.

Technical Skills

Languages: Python, C/C++, MATLAB, SQL, RISC-V, HiveQL

Tools/Frameworks: PyTorch, Transformers, LLMs, Google Cloud, Kubernetes, AWS S3, GitHub, Linux, Gradescope.